Solving large sparse linear systems with a variable s-step GMRES preconditioned by DD

David Imberti, Jocelyne Erhel

Solving large sparse linear systems with a variable s-step GMRES preconditioned by DD

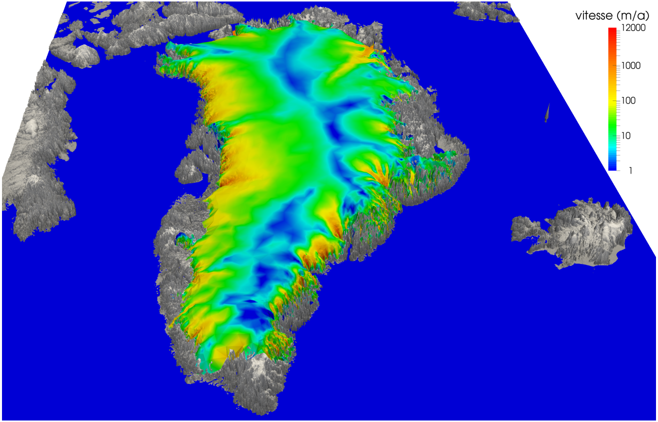
David Imberti and Jocelyne Erhel
Joint work with Désiré Nuentsa Wakam (first part)

FLUMINANCE team, Inria Rennes, France

DD24
Longyearbyen - Svalbard, Norway, February 2017

Numerical simulations

Solving large sparse linear systems is at the heart of many numerical simulations



Simulation of the velocity field on the Greenland inlandsis, with Elmer/Ice model
Image : Fabien Gillet-Chaulet, CNRS and LGGE, Grenoble.
More information on the website Interstices for the general public
["modeliser-et-simuler-la-fonte-des-calottes-polaires", Nodet and JE, 2015]

Outline

1. Krylov subspace algorithm GMRES
2. *m*-step preconditioned and deflated GMRES
3. Variable *s*-step algorithm VGMRES(m,s)

Preconditioned GMRES

$Ax = b, \quad A \in \mathbb{R}^{n\times n} \quad x, b \in \mathbb{R}^n \quad B = AM^{-1} \quad x_0 \in \mathbb{R}^n \quad r_0 = b - Ax_0$

GMRES(m): a Krylov subspace method

- [Saad and Schultz 1986, Meurant's book 1999, Saad's book 2003, Simoncini and Szyld 2007, JE 2011, ...]
- $\mathcal{K}_m(B, r_0) = span\{r_0, Br_0, \ldots B^{m-1}r_0\}$
- Find $x_m \in x_0 + M^{-1}\mathcal{K}_m(B, r_0)$ such that $\|r_m\|_2 = \|b - Bx_m\|_2 = \min_{x\in x_0+M^{-1}\mathcal{K}_m(B,r_0)} \|b - Bx\|_2$

Building blocks of GMRES(m)

- Build an orthonormal basis V_{k+1} of the Krylov subspace $\mathcal{K}_{k+1}$
  get the Arnoldi-like relation $BV_k = V_{k+1}H_k$ for $k = 1, \ldots, m$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_k y$ implies $r = r_0 - BV_k y = V_{k+1}(\beta e_1 - H_k y)$
  Solve the least-squares problem: $\min_{y\in\mathbb{R}^k} \|\beta e_1 - H_k y\|$
- Restart if not converged

$$x_0 = x_0 + M^{-1}V_m y_m$$

Preconditioned GMRES

$Ax = b, \quad A \in \mathbb{R}^{n \times n} \quad x, b \in \mathbb{R}^n \quad B = AM^{-1} \quad x_0 \in \mathbb{R}^n \quad r_0 = b - Ax_0$

GMRES(m): a Krylov subspace method

- [Saad and Schultz 1986, Meurant's book 1999, Saad's book 2003, Simoncini and Szyld 2007, JE 2011, ...]
- $\mathcal{K}_m(B, r_0) = span\{r_0, Br_0, \ldots B^{m-1} r_0\}$
- Find $x_m \in x_0 + M^{-1}\mathcal{K}_m(B, r_0)$ such that $\|r_m\|_2 = \|b - Bx_m\|_2 = \min_{x \in x_0 + M^{-1}\mathcal{K}_m(B, r_0)} \|b - Bx\|_2$

Building blocks of GMRES(m)

- Build an orthonormal basis V_{k+1} of the Krylov subspace $\mathcal{K}_{k+1}$
  get the Arnoldi-like relation $BV_k = V_{k+1}H_k$ for $k = 1, \ldots, m$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_k y$ implies $r = r_0 - BV_k y = V_{k+1}(\beta e_1 - H_k y)$
  Solve the least-squares problem: $\min_{y \in \mathbb{R}^k} \|\beta e_1 - H_k y\|$
- Restart if not converged

$$x_0 = x_0 + M^{-1} V_m y_m$$

GMRES ... practical issues

Arnoldi process

1: $v_1 = r_0/\|r_0\|_2$
2: **for** $k = 1, m$ **do**
3: $\quad p = Bv_k$
4: $\quad$ **for** $i = 1 : k$ **do**
5: $\quad\quad h_{ik} = v_i^T p$
6: $\quad\quad p = p - h_{ik} v_i$
7: $\quad$ **end for**
8: $\quad h_{k+1,k} = ||p||_2$
9: $\quad v_{k+1} = p/h_{k+1,k}$
10: **end for**

$\Downarrow$

$BV_m = V_{m+1} H_m$

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [JE 1995, Sidje 1997, Demmel et al 2011]
- Compute the basis block by block [De Sturler 1994, Hoemmen 2010]

Preconditioning issues

$\Rightarrow$ multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009,..]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ deflation to recover possible loss of information

- Deflation by preconditioning [JE et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam et al 2013, Nuentsa Wakam and Pacull 2013, Nuentsa Wakam and JE 2013]

GMRES ... practical issues

Arnoldi process

1: $v_1 = r_0 / \|r_0\|_2$
2: **for** $k = 1, m$ **do**
3: $p = Bv_k$
4: **for** $i = 1 : k$ **do**
5: $h_{ik} = v_i^T p$
6: $p = p - h_{ik} v_i$
7: **end for**
8: $h_{k+1,k} = ||p||_2$
9: $v_{k+1} = p / h_{k+1,k}$
10: **end for**

$\Downarrow$

$BV_m = V_{m+1} H_m$

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [JE 1995, Sidje 1997, Demmel et al 2011]
- Compute the basis block by block [De Sturler 1994, Hoemmen 2010]

Preconditioning issues

$\Rightarrow$ multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009,..]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ deflation to recover possible loss of information

- Deflation by preconditioning [JE et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners [Nuentsa Wakam et al 2013, Nuentsa Wakam and Pacull 2013, Nuentsa Wakam and JE 2013]

GMRES ... practical issues

Arnoldi process

```
v_1 = r_0/||r_0||_2
for k = 1, m do
    p = B v_k
    for i = 1 : k do
        h_ik = v_i^T p
        p = p - h_ik v_i
    end for
    h_{k+1,k} = ||p||_2
    v_{k+1} = p / h_{k+1,k}
end for
```

$\Downarrow$

$BV_m = V_{m+1} H_m$

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [JE 1995, Sidje 1997, Demmel et al 2011]
- Compute the basis block by block [De Sturler 1994, Hoemmen 2010]

Preconditioning issues

$\Rightarrow$ multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009,..]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ deflation to recover possible loss of information

- Deflation by preconditioning [JE et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam et al 2013, Nuentsa Wakam and Pacull 2013, Nuentsa Wakam and JE 2013]

GMRES ... practical issues

Arnoldi process

1: $v_1 = r_0/\|r_0\|_2$
2: **for** $k = 1, m$ **do**
3: $p = Bv_k$
4: **for** $i = 1 : k$ **do**
5: $h_{ik} = v_i^T p$
6: $p = p - h_{ik} v_i$
7: **end for**
8: $h_{k+1,k} = ||p||_2$
9: $v_{k+1} = p/h_{k+1,k}$
10: **end for**

$\Downarrow$

$BV_m = V_{m+1} H_m$

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [JE 1995, Sidje 1997, Demmel et al 2011]
- Compute the basis block by block [De Sturler 1994, Hoemmen 2010]

Preconditioning issues

$\Rightarrow$ multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009,..]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ deflation to recover possible loss of information

- Deflation by preconditioning [JE et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam et al 2013, Nuentsa Wakam and Pacull 2013, Nuentsa Wakam and JE 2013]

GMRES ... practical issues

Arnoldi process

$v_1 = r_0/\|r_0\|_2$
for $k = 1, m$ **do**
  $p = Bv_k$
  for $i = 1 : k$ **do**
    $h_{ik} = v_i^T p$
    $p = p - h_{ik} v_i$
  end for
  $h_{k+1,k} = ||p||_2$
  $v_{k+1} = p/h_{k+1,k}$
end for

$\Downarrow$

$BV_m = V_{m+1} H_m$

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [JE 1995, Sidje 1997, Demmel et al 2011]
- Compute the basis block by block [De Sturler 1994, Hoemmen 2010]

Preconditioning issues

$\Rightarrow$ multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009,..]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ deflation to recover possible loss of information

- Deflation by preconditioning [JE et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam et al 2013, Nuentsa Wakam and Pacull 2013, Nuentsa Wakam and JE 2013]

Outline

1. Krylov subspace algorithm GMRES
2. *m*-step preconditioned and deflated GMRES
3. Variable *s*-step algorithm VGMRES(m,s)

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

m-step GMRES(m)

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

m-step GMRES(m)

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

m-step GMRES(m)

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

m-step GMRES(m)

Building blocks of m-step GMRES(m)

- Build a basis W_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ such that $BW_m = W_{m+1}T_{m+1}$
- Build an orthonormal basis V_{m+1} of $\mathcal{K}_{m+1}$ such that $W_{m+1} = V_{m+1}R_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
  [Nuentsa Wakam and JE 2013, Hoemmen 2010]
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$
- Alternative to minimize the residual in the Krylov subspace
  [Imberti and JE 2016]
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:
  $$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$
- Restart if not converged

Choice of the basis W_m

Computation of vectors w_k

- Monomial basis: $w_{k+1} = Bw_k$
- Newton basis: $\sigma_{k+1} w_{k+1} = (B - \alpha_{k+1} I) w_k$
  where the shifts α_k are computed during a preliminary cycle of GMRES(m)
  and scaling is done with σ_k
  [Reichel 1990, Bai et al 1994, JE 1995, Nuentsa Wakam and JE 2013]
- Chebyshev basis [Joubert+Carey 1992, Philippe+Reichel 2012]

Stability and parallelism issues

- the condition number of W_m increases with m
- Parallel performances increases with m

GMRES combined with Domain Decomposition

Main steps with Domain Decomposition preconditioning

- Partition the weighted graph of the matrix in parallel with PARMETIS.
- Redistribute the matrix and right-hand-side according to the PARMETIS partitioning.
- Perform a parallel iterative row and column scaling on the matrix and the right-hand side vector [Amestoy et al 2008].
- Define the overlap between the submatrices for the additive Schwarz preconditioner [Cai and Sarkis 1999, Efstathiou and Gander 2003]

$$M_{RAS}^{-1} = \sum_{k=1}^{D} (R_k^0)^T (A_k^\delta)^{-1} R_k^\delta$$

- Setup the submatrices (ILU or LU factorization).
- Solve iteratively the preconditioned system using GMRES.

Deflation strategies

Restarted GMRES(m)

- The convergence rate depends on the spectral distribution in B
- Smallest eigenvalues slow down the convergence
- Deflation occurs when the Krylov subspace is large enough
- With restarting : loss of spectral information, risk of stalling

Accelerating the restarted GMRES [Simoncini and Szyld, 2007]

- Approximate the smallest eigenvalues and the associated invariant subspace U_r
- Explicit deflation technique
  [JE et al 1996; Burrage et al 1998; Moriya et al 2000, Nuentsa Wakam et al 2013]:

$$B\bar{M}^{-1}\tilde{x} = b$$

  with $\bar{M}^{-1} = (I_n + U_r(|\lambda_n|T^{-1} - I_r)U_r^T$ and $T = U_r^T B U_r$
- Augmented techniques
  [Morgan 2000, 2002, Giraud et al 2010, Nuentsa Wakam and JE 2013]

$$x_m \in x_0 + \mathcal{K}_m(B, r_0) + span\{U_r\}$$

Deflation strategies

Restarted GMRES(m)

- The convergence rate depends on the spectral distribution in B
- Smallest eigenvalues slow down the convergence
- Deflation occurs when the Krylov subspace is large enough
- With restarting : loss of spectral information, risk of stalling

Accelerating the restarted GMRES [Simoncini and Szyld, 2007]

- Approximate the smallest eigenvalues and the associated invariant subspace U_r
- Explicit deflation technique [JE et al 1996; Burrage et al 1998; Moriya et al 2000, Nuentsa Wakam et al 2013]:

$$B\bar{M}^{-1}\tilde{x} = b$$

with $\bar{M}^{-1} = (I_n + U_r(|\lambda_n|T^{-1} - I_r)U_r^T$ and $T = U_r^T B U_r$

- Augmented techniques [Morgan 2000, 2002, Giraud et al 2010, Nuentsa Wakam and JE 2013]

$$x_m \in x_0 + \mathcal{K}_m(B, r_0) + span\{U_r\}$$

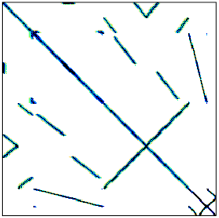
Experiments with CFD matrices

FLUOREM matrices

- in MatrixMarket collection
- large, sparse, nonsymmetric matrices
- linearization of Navier-Stokes: symmetric profile with structured blocks

[Nuentsa Wakam+Pacull 2013]

RM07R n= 381,689; nnz=37,464,962



Software

Krylov solvers integrated in Petsc
[Nuentsa Wakam 2011]

- Schwarz preconditioning combined with GMRES, Newton basis and deflation
- DGMRES: preconditioning deflation
- AGMRES: augmented subspace

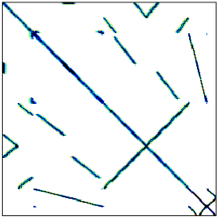
Experiments with CFD matrices

FLUOREM matrices

- in MatrixMarket collection
- large, sparse, nonsymmetric matrices
- linearization of Navier-Stokes: symmetric profile with structured blocks

[Nuentsa Wakam+Pacull 2013]

RM07R n= 381,689; nnz=37,464,962



Software

Krylov solvers integrated in Petsc
[Nuentsa Wakam 2011]

- Schwarz preconditioning combined with GMRES, Newton basis and deflation
- DGMRES: preconditioning deflation
- AGMRES: augmented subspace

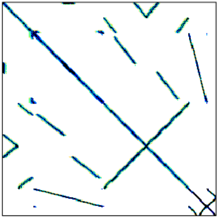
Experiments with CFD matrices

FLUOREM matrices

- in MatrixMarket collection
- large, sparse, nonsymmetric matrices
- linearization of Navier-Stokes: symmetric profile with structured blocks

[Nuentsa Wakam+Pacull 2013]

RM07R n= 381,689; nnz=37,464,962



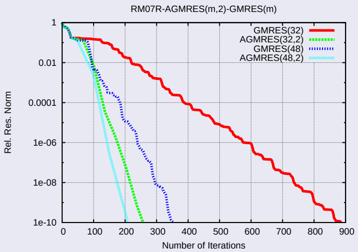
Software

Krylov solvers integrated in Petsc
[Nuentsa Wakam 2011]

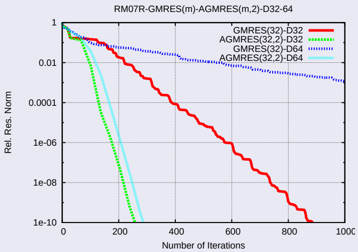
- Schwarz preconditioning combined with GMRES, Newton basis and deflation
- DGMRES: preconditioning deflation
- AGMRES: augmented subspace

Convergence with Augmented GMRES (AGMRES)

RM07R: effect of the restarting



RM07R: effect of the number of subdomains



CPU time on parallel computers

RM07R, n = 381,689, nz = 37,464,962				
D	GMRES(32)		AGMRES(32,r)	
	ITS	**Time (s)**	**ITS**	**Time (s)**
16	254	379.3	169	224.1
32	886	573.4	212	91.41
64	-	-	287	62.39

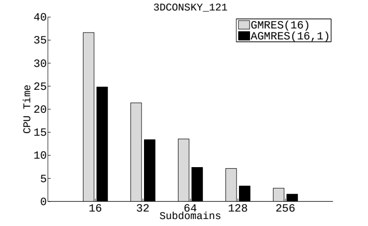
Parallel CPU Time with AGMRES

Convection-Diffusion test cases

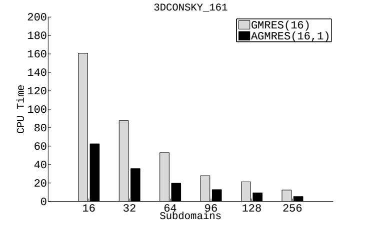
- 3DCONSKY_121 : size = 1,771,561; nonzeros = 50,178,241
- 3DCOSKY_161 : size= 4,173,281; nonzeros = 118,645,121

[Nuentsa Wakam and JE 2013]

3DCONSKY_121



3DCONSKY_161



Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1$, $l_j = l_{j-1} + s_j$, $j \geq 2$
[Imberti and JE 2016]

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s-step VGMRES(m,s)

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$
[Imberti and JE 2016]

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s-step VGMRES(m,s)

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$
[Imberti and JE 2016]

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s-step VGMRES(m,s)

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$
[Imberti and JE 2016]

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s-step VGMRES(m,s)

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$
[Imberti and JE 2016]

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1} W_m y_m$$

Block computation and orthogonalization with VGMRES(m,s)

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0$, $\beta = \|r_0\|$, $v_1 = r_0/\beta$, $W_0 = \emptyset$, $V_1 = [v_1]$

Block j of size s_j, with $1 \leq j \leq J$
with an adaptive choice of s_j such that $s_j \leq s$

- First step: s_j matrix vector products
  Compute the block C_j as a basis of $\mathcal{K}_{s_j}(B, u)$ with $u = v_{l_{j-1}+1}$
  Compute the block BC_j
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Define the Krylov basis by

$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{l_j+1} S_j$$

RODDEC [Sidje 1997, JE 1995] or TSQR [Demmel et al 2011]
By induction, get the Arnoldi-like relation

$$\left[v_1, BW_{l_j}\right] = V_{l_j+1} R_{l_j+1}, \; BW_{l_j} = V_{l_j+1} H_{l_j}$$

Block computation and orthogonalization with VGMRES(m,s)

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$
with an adaptive choice of s_j such that $s_j \leq s$

- First step: s_j matrix vector products
  Compute the block C_j as a basis of $\mathcal{K}_{s_j}(B, u)$ with $u = v_{l_{j-1}+1}$
  Compute the block BC_j
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Define the Krylov basis by
$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$
- Second step: orthogonalization
$$BC_j = V_{l_j+1} S_j$$
  RODDEC [Sidje 1997, JE 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation
$$\left[v_1, BW_{l_j}\right] = V_{l_j+1} R_{l_j+1}, BW_{l_j} = V_{l_j+1} H_{l_j}$$

Block computation and orthogonalization with VGMRES(m,s)

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$
with an adaptive choice of s_j such that $s_j \leq s$

- First step: s_j matrix vector products
  Compute the block C_j as a basis of $\mathcal{K}_{s_j}(B, u)$ with $u = v_{I_{j-1}+1}$
  Compute the block BC_j
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Define the Krylov basis by
  $$W_{I_j} = \left[W_{I_{j-1}}, C_j\right]$$
- Second step: orthogonalization
  $$BC_j = V_{I_j+1}S_j$$
  RODDEC [Sidje 1997, JE 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation
  $$\left[v_1, BW_{I_j}\right] = V_{I_j+1}R_{I_j+1}, BW_{I_j} = V_{I_j+1}H_{I_j}$$

Block computation and orthogonalization with VGMRES(m,s)

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$
with an adaptive choice of s_j such that $s_j \leq s$

- First step: s_j matrix vector products
  Compute the block C_j as a basis of $\mathcal{K}_{s_j}(B, u)$ with $u = v_{l_{j-1}+1}$
  Compute the block BC_j
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Define the Krylov basis by
  $$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$
- Second step: orthogonalization
  $$BC_j = V_{l_j+1} S_j$$
  RODDEC [Sidje 1997, JE 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation
  $$\left[v_1, BW_{l_j}\right] = V_{l_j+1} R_{l_j+1}, BW_{l_j} = V_{l_j+1} H_{l_j}$$

Condition number of BC_j and BW_{I_j}

- block basis

$$W_{I_j} = [C_1, C_2, \ldots, C_j]$$

$$\kappa(BW_{I_j}) \geq \max_{1 \leq k \leq j} \kappa(BC_k)$$

- symmetric case: exponential growth of the condition number of a Krylov basis
  [Beckermann 2000]
- nonsymmetric case with $|\lambda_1| > |\lambda_2|$
  monomial basis $C_j = \{u, Bu, \ldots, B^{s_j - 1}u\}$
  [Imberti and JE 2016]

$$\kappa(BC_j) \geq cste\, |\lambda_1/\lambda_2|^{s_j - 1}$$

Choice of the block size s_j

Objective: small condition numbers of the first blocks

- fixed sequence $s_j = s$: SGMRES(m,s)
  [Hoemmen 2010, ...]
- adaptive increasing sequence s_j with $s_j \leq s$: VGMRES(m,s)
  [Imberti and JE 2016]

Stability issues

Condition number of BC_j and BW_{I_j}

- block basis

$$W_{I_j} = [C_1, C_2, \dots, C_j]$$

$$\kappa(BW_{I_j}) \geq \max_{1 \leq k \leq j} \kappa(BC_k)$$

- symmetric case: exponential growth of the condition number of a Krylov basis [Beckermann 2000]
- nonsymmetric case with $|\lambda_1| > |\lambda_2|$
  monomial basis $C_j = \{u, Bu, \dots, B^{s_j - 1} u\}$
  [Imberti and JE 2016]

$$\kappa(BC_j) \geq cste\, |\lambda_1/\lambda_2|^{s_j - 1}$$

Choice of the block size s_j

Objective: small condition numbers of the first blocks

- fixed sequence $s_j = s$: SGMRES(m,s)
  [Hoemmen 2010, ...]
- adaptive increasing sequence s_j with $s_j \leq s$: VGMRES(m,s)
  [Imberti and JE 2016]

Communication issues

The number of messages is related to the number of steps J
Objective: reduce the number of steps

Number of steps

- **SGMRES(m,s):** $J = m/s$ steps
- **VGMRES(m,s):** J steps with $I_J = m$ and $J = J_1 + J_2$
- **FibGMRES(m,s):** Fibonacci sequence capped at s

$$J_2 = O(m/s),\ J_1 = O(\log_\phi(s))$$

The number of messages is related to the number of steps J
Objective: reduce the number of steps

Number of steps

- **SGMRES(m,s):** $J = m/s$ steps
- **VGMRES(m,s):** J steps with $l_J = m$ and $J = J_1 + J_2$
- **FibGMRES(m,s):** Fibonacci sequence capped at s

$$J_2 = O(m/s), J_1 = O(\log_\phi(s))$$

Sequences with $m = 48$ and $s = 16$

j	1	2	3	4	5	6	7
s_j	1	2	3	5	8	13	16
l_j	1	3	6	11	19	32	48

Variable increasing block size for $m = 48$ **and** $s = 16$**.**

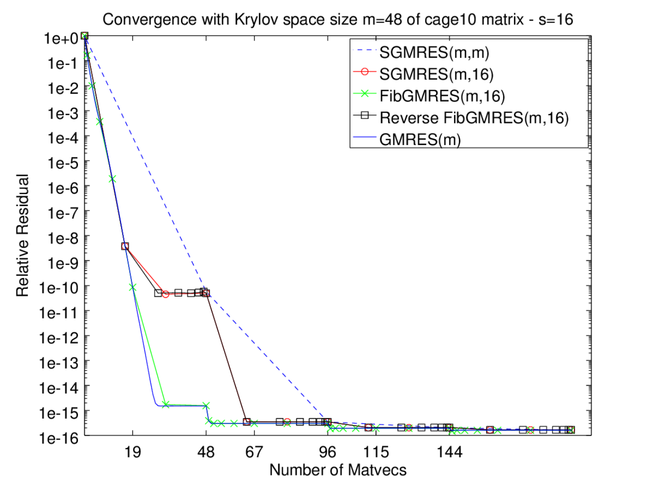
j	1	2	3	4	5	6	7
s_j	16	13	8	5	3	2	1
l_j	16	29	37	42	45	47	48

Variable decreasing block size for $m = 48$ **and** $s = 16$**.**

Numerical experiments done with a monomial basis

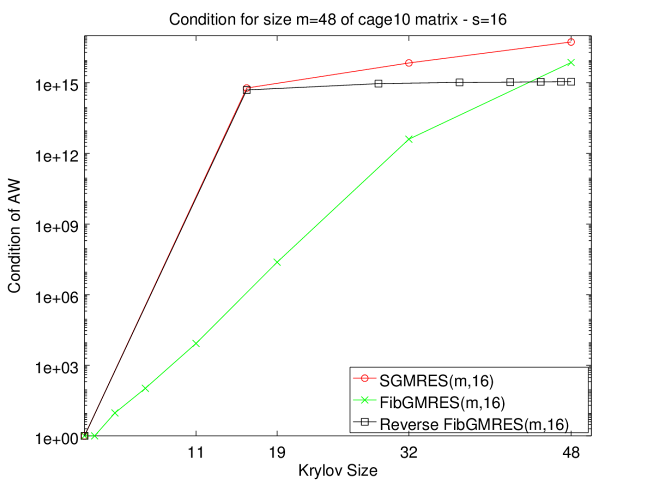
Nonsymmetric matrix CAGE10 of size $n = 11397$ and nonzeros $nz = 150645$

Convergence curves with $m = 48$ and $s = 16$



Nonsymmetric matrix CAGE10 of size $n = 11397$ and nonzeros $nz = 150645$

Condition numbers of AW with $m = 48$ and $s = 16$



Sequences of FibGMRES(m,s) with $m = 96$

j	1	2	3	4	5	6	7	8	9	10
s_j	1	2	3	5	8	13	16	16	16	16
l_j	1	3	6	11	19	32	48	64	80	96

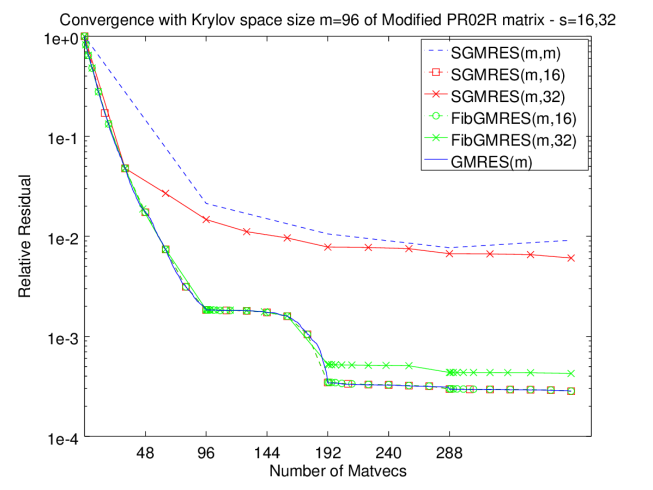
Variable increasing block size for $m = 96$ **and** $s = 16$

j	1	2	3	4	5	6	7	8	9
s_j	1	2	3	5	8	13	14	18	32
l_j	1	3	6	11	19	32	46	64	96

Variable increasing block size for $m = 96$ **and** $s = 32$

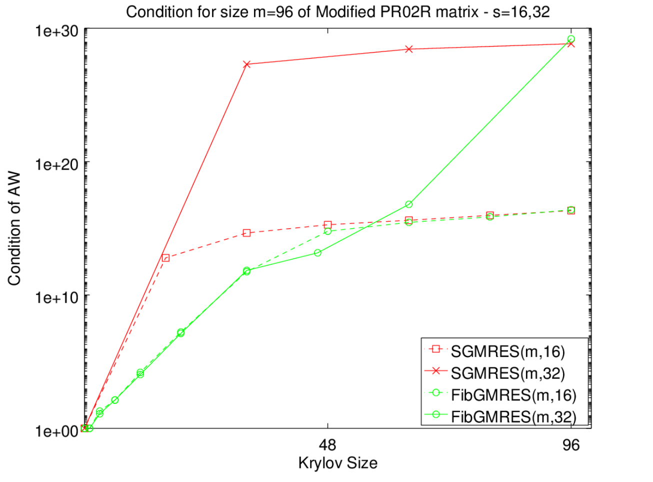
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Convergence curves with $m = 96$ and $s = 16$ or $s = 32$



Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

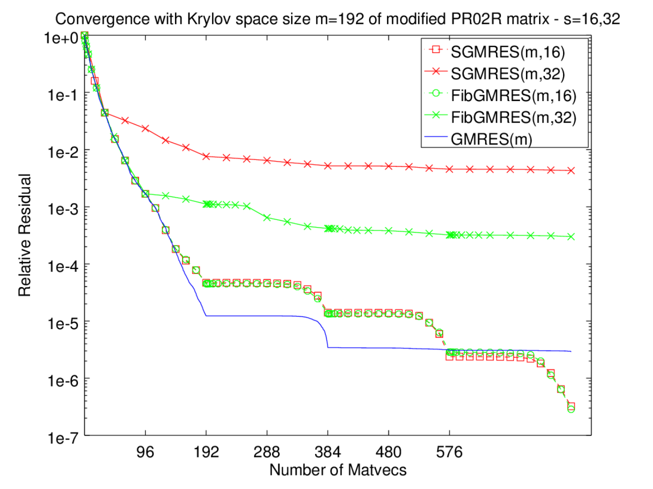
Condition numbers of AW with $m = 96$ and $s = 16$ or $s = 32$



Numerical experiment with a large restarting parameter

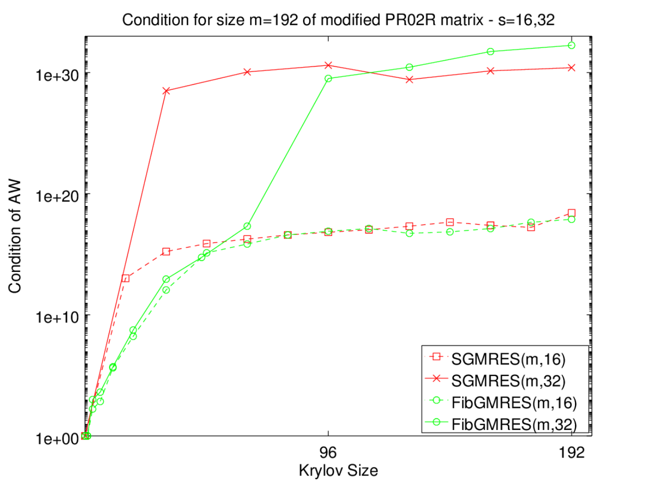
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Convergence curves with $m = 192$ **and** $s = 16$ **or** $s = 32$



Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Condition numbers of AW with $m = 192$ and $s = 16$ or $s = 32$



Conclusion

DGMRES(m,r) and AGMRES(m,r)

- Combined DD preconditioning with m-step GMRES and deflation
- Parallel efficiency and fast convergence with a large number of subdomains
- Difficult choice of the restarting parameter m

VGMRES(m,s) and FibGMRES(m,s)

- s-step algorithm combined with a variable block size s_j
- Relationship between convergence rate and condition numbers of the blocks
- Numerical experiments with a Fibonacci sequence

Future work

- Adaptive variation of s
- Blocks computed via a Newton basis
- Schwarz preconditioning with deflation
- Parallel computations

Conclusion

DGMRES(m,r) and AGMRES(m,r)

- Combined DD preconditioning with m-step GMRES and deflation
- Parallel efficiency and fast convergence with a large number of subdomains
- Difficult choice of the restarting parameter m

VGMRES(m,s) and FibGMRES(m,s)

- s-step algorithm combined with a variable block size s_j
- Relationship between convergence rate and condition numbers of the blocks
- Numerical experiments with a Fibonacci sequence

Future work

- Adaptive variation of s
- Blocks computed via a Newton basis
- Schwarz preconditioning with deflation
- Parallel computations

Conclusion

DGMRES(m,r) and AGMRES(m,r)

- Combined DD preconditioning with m-step GMRES and deflation
- Parallel efficiency and fast convergence with a large number of subdomains
- Difficult choice of the restarting parameter m

VGMRES(m,s) and FibGMRES(m,s)

- s-step algorithm combined with a variable block size s_j
- Relationship between convergence rate and condition numbers of the blocks
- Numerical experiments with a Fibonacci sequence

Future work

- Adaptive variation of s
- Blocks computed via a Newton basis
- Schwarz preconditioning with deflation
- Parallel computations